J

15

20

25

35

SEQUENCE LISTING

The following Sequence Listing is submitted pursuant to 37 CFR §1.821. A copy in computer readable form is also submitted herewith. The paper and computer readable forms of this Sequence Listing are the same.

1.0130

(1) GENERAL INFORMATION:

- (i) APPLICANT: Tripp, Cynthia Ann Frank, Glenn R. Grieve, Robert B.
- 10 (ii) TITLE OF INVENTION: NOVEL PARASITE ASTACIN METALLOENDOPEPTIDASE PROTEINS
 - (iii) NUMBER OF SEQUENCES: 36
 - (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: SHERIDAN ROSS P.C.
 - (B) STREET: 1700 LINCOLN ST., SUITE 3500
 - (C) CITY: DENVER
 - (D) STATE: CO
 - (E) COUNTRY: USA
 - (F) ZIP: 80203
 - (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 - (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- 30 (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Connell, Gary J.
 - (B) REGISTRATION NUMBER: 32,020
 - (C) REFERENCE/DOCKET NUMBER: 2618-21-1-C1
 - (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (303) 863-9700
 - (B) TELEFAX: (303) 863-0223

10

15

20

25

30

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1299 base pairs

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

$\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}\mathbf{T}$	TTTTTTTTGT	TTCATTGTTC	AGTCAGTGGA	AAATTATCGA	ACGCAGAAAG	60
CATCACGAAA	TACGTTAGAT	CACATCAAAC	AACTTATCAC	CTTGAACGTA	CAAAGAGAGA	120
TTGGAAACAT	AGATGATAAG	ACATTAGCTG	ATGAAATAGT	ATTACAACGA	CGGGATCCTG	180
AGGCAAAATG	GCATCATAAT	GAACTATTCA	TTAATGATCC	AGATGCATAC	TATCAAGGCG	240
ATGTCGATTT	GTCGGAAAAA	CAAGCCGAAA	TTCTAAGCGA	ACATTTTAAA	AATGAAATTG	300
CTTTAACAGA	GAAAGACGAC	ACAATAATAC	GGCGAAAAAA	GAGCATTGGT	CGTGAACCAT	360
TTTACGTAAG	ATGGAATCAT	AAACGTCCCA	TTAGCTATGA	ATTTGCGGAA	AGTATTCCAT	420
TAGAAACACG	TAGAAAAATT	CGTTCAGCAA	TAGCAATGTG	GGAAGAACGA	ACATGCATAC	480
GATTCCAAGA	AAATGGCCCA	AATGTAGATC	GAATTGAATT	TTACGACGGT	GGCGGTTGTT	540
CAAGTTTTGT	CGGCCGAACA	GGAGGGAATT	TCAATTTCAA	CACCAGGATG	TGATATTATT	600
GGTATTATAT	CACATGAAAT	TGGTCATACT	TTAGGAATAT	TTCATGAGCA	AGCACGTCGT	660
GATCAAAAAA	ATCATATTTT	TATTAATTAC	AACAATATTC	CATCAAGCCG	TTGGAACAAT	720
TTTTTTCCAT	TATCAGAATA	TGAAGCTGAT	ATGTTTAATT	TACCTTATGA	TACAGGATCA	780
GTAATGCACT	ATGGTTCATA	CGGATTTGCA	AGAAATCCGT	ATGAACCAAC	TATTACAACA	840
CGTGATAAAT	TTCAACAGTA	CACAATTGGG	CAACGTGAAG	GGCCATCATT	TCTGGATTAT	900
GCATCTGTTA	AGCTTTATCT	ACAAACGCAT	TAATGATATT	GTTATCAAAT	GGATGATAAT	960
TTCAATAAGT	ATAAACAGCG	CTTATCGTTG	TACAGAACAA	TGTGCTGATA	TGCACTGCGA	1020
TCATAATGGT	TATCCGGATC	CTAATAATTG	CGCGAAATGC	TTGTGTCCAG	ATGGTTTTGC	1080
TGGTCGTACC	TGTCAATTTG	TTCAATATAC	ATCTTGCGGA	GCTCTCATTA	AGGTAAGTAT	1140
TGTCTTTTGA	CCTCTTCTCT	GACTAAAATA	TAAGTTAAGC	ATATGTATCT	TCCGTCTAAT	1200
GATTTTCTTG	ATTTTGATTT	GTTCAATGCT	CTTCTTGATA	ATAATATAAA	AATTTTTGAA	1260
AATAAAGTTA	ACTTTTGGTC	ААААААААА	AAAAAAAA			1299

10

15

20

25

30

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2126 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GAAAGCATCA	CGAAATACGT	TAGATCACAT	CAAACAACTT	ATCACCTTGA	ACGTACAAAG	60
AGAGATTGGA	AACATAGATG	ATAAGACATT	AGCTGATGAA	ATAGTATTAC	AACGACGGGA	120
TCCTGAGGCA	AAATGGCATC	ATAATGAACT	ATTCATTAAT	GATCCAGATG	CATACTATCA	180
AGGCGATGTC	GATTTGTCGG	AAAAACAAGC	CGAAATTCTA	AGCGAACATT	TTAAAAATGA	240
AATTGCTTTA	ACAGAGAAAG	ACGACACAAT	AATACGGCGA	AAAAAGAGCA	TTGGTCGTGA	300
ACCATTTTAC	GTAAGATGGA	ATCATAAACG	TCCCATTAGC	TATGAATTTG	CGGAAAGTAT	360
TCCATTAGAA	ACACGTAGAA	AAATTCGTTC	AGCAATAGCA	ATGTGGGAAG	AACGAACATG	420
CATACGATTC	CAAGAAAATG	GCCCAAATGT	AGATCGAATT	GAATTTTACG	ACGGTGGCGG	480
TTGTTCAAGT	TTTGTCGGCC	GAACAGGAGG	GAATTTCAAT	TTCAACACCA	GGATGTGATA	540
TTATTGGTAT	TATATCACAT	GAAATTGGTC	ATACTTTAGG	AATATTTCAT	GAGCAAGCAC	600
GTCGTGATCA	AAAAAATCAT	ATTTTTTATTA	ATTACAACAA	TATTCCATCA	AGCCGTTGGA	660
ACAATTTTTT	TCCATTATCA	GAATATGAAG	CTGATATGTT	TAATTTACCT	TATGATACAG	720
GATCAGTAAT	GCACTATGGT	TCATACGGAT	TTGCAAGAAA	TCCGTATGAA	CCAACTATTA	780
CAACACGTGA	TAAATTTCAA	CAGTACACAA	TTGGGCAACG	TGAAGGGCCA	TCATTTCTGG	840
ATTATGCATC	TGATAAACAG	CGCTTATCGT	TGTACAGAAC	AATGTGCTGA	TATGCACTGC	900
GATCATAATG	GTTATCCGGA	тсстаатаат	TGCGCGAAAT	GCTTGTGTCC	AGATGGTTTT	960
GCTGGTCGTA	CCTGTCAATT	TGTTCAATAT	ACATCTTGCG	GAGCTCTCAT	TAAGGCGAGG	1020
AAAATGCCTG	TTACGATTTC	GAGCCCAAAT	TATCCAAACT	TCTTCAATGT	TGGTGATCAA	1080
TGTATTTGGT	TGCTTACAGC	TCCACGCGTG	ATTCGTAAAT	TTGCAGTTTG	TTGAACAATT	1140
TCAATTACAA	TGTGAAGATA	CGTGTGATAA	ATCCTATGTA	GAAGTGAAAG	CTGACGCTGA	1200
TTTTCGACCT	ACTGGATATC	GATTTTGTTG	TTCGCGAGTG	CCACGTCATA	TTTTTCAATC	1260
TGCGACAAAC	GAGATGGTAG	TAATATTTCG	CGGTTTTGGT	GATGCGGGAA	ATGGCTTTAA	1320

10

20

25

30

AGCTAAAATT	TGGTCAAACG	TAGATGATGA	TATAGCTAAT	ACAATTGTAA	CAACTGAAAT	1380
GGCAAAAATT	TCGGAAAAAA	TACCGAAGCT	AACAGTTCCA	ATAGTTAAAA	CTATTACCAC	1440
TCCTACAATA	ACAACTACTA	CTGCTTTCAT	GATATCACCC	AAGAAAGGCA	ATGTCACCGC	1500
CACGAGAGTT	GCTATCACTA	CTACGCCGAC	TACTACAATT	ACTACGACTA	TTGCCGGTAC	1560
GTACCAATCA	CCGTAACTAA	TAATACTACA	CCTGTAGTAA	GTGAAACTTT	ACCATCATTG	1620
CCAGTCAAGA	TTCGAAACAA	AATAGGTGCA	TGCGAATGTG	GTGAATGGAC	AGAATGGACA	1680
GGTCCATGCT	CTCAAGAATG	TGGCGGTTGC	GGAAAACGTC	TTCGAACACG	TCAGTGTTCA	1740
TCAGATACGG	AATGTAGAAC	AGAAGAAAA	CGTGCGTGTG	CTTTTAAGTT	TGCCCATACG	1800
GGACTAATTT	CCTTATCAAT	AATGGAGAGT	TTCATATACT	TTGGAAGGGC	TGCTGTGTTG	1860
GTCTATTCCG	ATCGGGAGAT	ATGTGTTCAG	CACTTGATGA	TAACGAGAAT	CCATTTCTGA	1920
AATTTCTAGA	ATCACTGTTG	AACATGCAAG	ATTCTCGAAA	AAACGATAAT	TTGCCTGACT	1980
CGAAAAAGAA	GTGATTGAAT	GATTCGATAA	TATTGATTAA	TAAAACGGGT	TGTATTCTCG	2040
TCATAGAGTA	TCCGTTGATG	TTTTTATCCA	AAAAATTCTC	TTGCTTTTAA	TTATTGTGAA	2100
TAAAACTTTT	GTTTACCCAA	AAAAA				2126

15 (2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 191 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Arg Asn Thr Leu Asp His Ile Lys Gln Leu Ile Thr Leu Asn Val Gln
20 25 30

Arg Glu Ile Gly Asn Ile Asp Asp Lys Thr Leu Ala Asp Glu Ile Val 35 40 45

Leu Gln Arg Arg Asp Pro Glu Ala Lys Trp His His Asn Glu Leu Phe 50 55 60

Ile Asn Asp Pro Asp Ala Tyr Tyr Gln Gly Asp Val Asp Leu Ser Glu 65 70 75 80

		гур	GTII	AIa	GIU	85	ьец	ser	GIU	HIS	90	ьys	Asn	GIU	lle	95	ьeu
		Thr	Glu	Lys	Asp 100	Asp	Thr	Ile	Ile	Arg 105	Arg	Lys	Lys	Ser	Ile 110	Gly	Arg
5		Glu	Pro	Phe 115	Tyr	Val	Arg	Trp	Asn 120	His	Lys	Arg	Pro	Ile 125	Ser	Tyr	Glu
		Phe	Ala 130	Glu	Ser	Ile	Pro	Leu 135	Glu	Thr	Arg	Arg	Lys 140	Ile	Arg	Ser	Ala
10		Ile 145	Ala	Met	Trp	Glu	Glu 150	Arg	Thr	Cys	Ile	Arg 155	Phe	Gln	Glu	Asn	Gly 160
		Pro	Asn	Val	Asp	Arg 165	Ile	Glu	Phe	Tyr	Asp 170	Gly	Gly	Gly	Cys	Ser 175	Ser
		Phe	Val	Gly	Arg 180	Thr	Gly	Gly	Asn	Phe 185	Asn	Phe	Asn	Thr	Arg 190	Met	
15	(2) 1	INFO	RMATI	ION I	FOR S	SEQ :	ID N	0:4:									
		(i)	(A) (B) (C)	LEN TYI	E CHANGTH: PE: &	: 141 amino EDNES	l am: o ac: SS:	ino a id		5							
20			(D)	TOI	OLOG	GY: I	linea	ar									
	((ii)	MOLE	ECULE	TYI	PE: p	prote	ein									
	((xi)	SEQU	JENCE	E DES	SCRI	OITS	V: SE	EQ II	ON C	: 4:						
		Ile 1	Glu	Leu	Asn	Phe 5	Thr	Thr	Val	Ala	Val 10	Val	Gln	Val	Leu	Ser 15	Ala
25		Glu	Gln	Glu	Gly 20	Ile	Ser	Ile	Ser	Thr 25	Pro	Gly	Cys	Asp	Ile 30	Ile	Gly
		Ile	Ile	Ser 35	His	Glu	Ile	Gly	His 40	Thr	Leu	Gly	Ile	Phe 45	His	Glu	Gln
30		Ala	Arg 50	Arg	Asp	Gln	Lys	Asn 55	His	Ile	Phe	Ile	Asn 60	Tyr	Asn	Asn	Ile
		Pro 65	Ser	Ser	Arg	Trp	Asn 70	Asn	Phe	Phe	Pro	Leu 75	Ser	Glu	Tyr	Glu	Ala 80
		Asp	Met	Phe	Asn	Leu 85	Pro	Tyr	Asp	Thr	Gly 90	Ser	Val	Met	His	Tyr 95	Gly
35		Ser	Tyr	Gly	Phe 100	Ala	Arg	Asn	Pro	Туг 105	Glu	Pro	Thr	Ile	Thr 110	Thr	Arg
		Asp	Lys	Phe 115	Gln	Gln	Tyr	Thr	Ile 120	Gly	Gln	Arg	Glu	Gly 125	Pro	Ser	Phe

30

5

Leu Asp Tyr Ala Ser Val Lys Leu Tyr Leu Gln Thr His 130 135 140

- (2) INFORMATION FOR SEQ ID NO:5:
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 121 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Cys Thr Met Val His Thr Asp Leu Gln Glu Ile Arg Met Asn Gln Leu
1 5 10 15

Leu Gln His Val Ile Asn Phe Asn Ser Thr Gln Leu Gly Asn Val Lys 20 25 30

Gly His His Phe Trp Ile Met His Leu Leu Ser Phe Ile Tyr Lys Arg
35 40 45

Ile Asn Asp Ile Val Ile Lys Trp Met Ile Ile Ser Ile Ser Ile Asn 50 55 60

Ser Ala Tyr Arg Cys Thr Glu Gln Cys Ala Asp Met His Cys Asp His 65 70 75 80

Asn Gly Tyr Pro Asp Pro Asn Asn Cys Ala Lys Cys Leu Cys Pro Asp 85 90 95

Gly Phe Ala Gly Arg Thr Cys Gln Phe Val Gln Tyr Thr Ser Cys Gly

- Ala Leu Ile Lys Val Ser Ile Val Phe
 - (2) INFORMATION FOR SEQ ID NO:6:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 178 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Lys Ala Ser Arg Asn Thr Leu Asp His Ile Lys Gln Leu Ile Thr Leu
1 5 10 15

Asn Val Gln Arg Glu Ile Gly Asn Ile Asp Asp Lys Thr Leu Ala Asp 20 25 30

		Glu	Ile	Val 35	Leu	Gln	Arg	Arg	Asp 40	Pro	Glu	Ala	Lys	Trp 45	His	His	Asn
		Glu	Leu 50	Phe	Ile	Asn	Asp	Pro 55	Asp	Ala	Tyr	Tyr	Gln 60	Gly	Asp	Val	Asp
5		Leu 65	Ser	Glu	Lys	Gln	Ala 70	Glu	Ile	Leu	Ser	Glu 75	His	Phe	Lys	Asn	Glu 80
٠		Ile	Ala	Leu	Thr	Glu 85	Lys	Asp	Asp	Thr	Ile 90	Ile	Arg	Arg	Lys	Lys 95	Ser
10		Ile	Gly	Arg	Glu 100	Pro	Phe	Tyr	Val	Arg 105	Trp	Asn	His	Lys	Arg 110	Pro	Ile
		Ser	Tyr	Glu 115	Phe	Ala	Glu	Ser	Ile 120	Pro	Leu	Glu	Thr	Arg 125	Arg	Lys	Ile
		Arg	Ser 130	Ala	Ile	Ala	Met	Trp 135	Glu	Glu	Arg	Thr	Cys 140	Ile	Arg	Phe	Gln
15		Glu 145	Asn	Gly	Pro	Asn	Val 150	Asp	Arg	Ile	Glu	Phe 155	Tyr	Asp	Gly	Gly	Gly 160
		Cys	Ser	Ser	Phe	Val 165	Gly	Arg	Thr	Gly	Gly 170	Asn	Phe	Asn	Phe	Asn 175	Thr
20		Arg	Met														
	(2)	INFO	RMATI	ON I	OR S	SEQ 1	D NO):7:									
25		(i)	(A) (B)	STF	NGTH: PE: a	: 145 amino EDNES	ami aci SS:	ino a		5							
		(ii)	MOLE	CULE	TY!	PE: p	rote	ein									
		(xi)	SEQU	JENCE	E DES	CRIE	OITS	1: SE	EQ II	ONO:	7:						
30		Ile 1	Glu	Leu	Asn	Phe 5	Thr	Thr	Val	Ala	Val 10	Val	Gln	Val	Leu	Ser 15	Ala
		Glu	Gln	Glu	Gly 20	Ile	Ser	Ile	Ser	Thr 25	Pro	Gly	Cys	Asp	Ile 30	Ile	Gly
		Ile	Ile	Ser 35	His	Glu	Ile	Gly	His 40	Thr	Leu	Gly	Ile	Phe 45	His	Glu	Gln
35		Ala	Arg 50	Arg	Asp	Gln	Lys	Asn 55	His	Ile	Phe	Ile	Asn 60	Tyr	Asn	Asn	Ile
		Pro 65	Ser	Ser	Arg	Trp	Asn 70	Asn	Phe	Phe	Pro	Leu 75	Ser	Glu	Tyr	Glu	Ala 80

	Asp	Met	Phe	Asn	Leu 85	Pro	Tyr	Asp	Thr	Gly 90	Ser	Val	Met	His	Tyr 95	Gly
	Ser	Tyr	Gly	Phe 100	Ala	Arg	Asn	Pro	Tyr 105	Glu	Pro	Thr	Ile	Thr 110	Thr	Arg
5	Asp	Lys	Phe 115	Gln	Gln	Tyr	Thr	Ile 120	Gly	Gln	Arg	Glu	Gly 125	Pro	Ser	Phe
	Leu	Asp 130	Tyr	Ala	Ser	Asp	Lys 135	Gln	Arg	Leu	Ser	Leu 140	Tyr	Arg	Thr	Met
10	Cys 145															
	(2) INFO	RMATI	ON I	OR S	SEQ 3	ID NO	0:8:									
15	(i)	(B) (C)	LEN TYI STI	NGTH: PE: & RANDE		4 ami o aci SS:	ino a id		5							
	(ii)	MOLE	ECULE	TYI	?E: p	prote	ein									
	(xi)	SEQU	JENCI	E DES	SCRI	OITS	N: SE	EQ II	ONO:	:8:						
20	Cys 1	Thr	Met	Val	His 5	Thr	Asp	Leu	Gln	Glu 10	Ile	Arg	Met	Asn	Gln 15	Leu
	Leu	Gln	His	Val 20	Ile	Asn	Phe	Asn	Ser 25	Thr	Gln	Leu	Gly	Asn 30	Val	Lys
	Gly	His	His 35	Phe	Trp	Ile	Met	His 40	Leu	Ile	Asn	Ser	Ala 45	Tyr	Arg	Cys
25	Thr	Glu 50	Gln	Cys	Ala	Asp	Met 55	His	Cys	Asp	His	Asn 60	Gly	Tyr	Pro	Asp
	Pro 65	Asn	Asn	Cys	Ala	Lys 70	Cys	Leu	Cys	Pro	Asp 75	Gly	Phe	Ala	Gly	Arg 80
30	Thr	Cys	Gln	Phe	Val 85	Gln	Tyr	Thr	Ser	Cys 90	Gly	Ala	Leu	Ile	Lys 95	Ala
	Arg	Lys	Met	Pro 100	Val	Thr	Ile	Ser	Ser 105	Pro	Asn	Tyr	Pro	Asn 110	Phe	Phe
	Asn	Val	Gly 115	Asp	Gln	Cys	Ile	Trp 120	Leu	Leu	Thr	Ala	Pro 125	Arg	Val	Ile
35	Arg	Lys 130	Phe	Ala	Val	Cys										

5

(2) INFORMATION FOR SEQ ID NO:9: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 154 amino acids (B) TYPE: amino acid (C) STRANDEDNESS:

- (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:
- Phe Val Asn Leu Gln Phe Val Glu Gln Phe Gln Leu Gln Cys Glu Asp
 10 1 5 10 15

Thr Cys Asp Lys Ser Tyr Val Glu Val Lys Ala Asp Ala Asp Phe Arg 20 25 30

Pro Thr Gly Tyr Arg Phe Cys Cys Ser Arg Val Pro Arg His Ile Phe 35 40 45

Ala Gly Asn Gly Phe Lys Ala Lys Ile Trp Ser Asn Val Asp Asp Asp 65 70 75 80

Ile Ala Asn Thr Ile Val Thr Thr Glu Met Ala Lys Ile Ser Glu Lys 85 90 95

Ile Pro Lys Leu Thr Val Pro Ile Val Lys Thr Ile Thr Thr Pro Thr 100 105 110

Ile Thr Thr Thr Thr Ala Phe Met Ile Ser Pro Lys Lys Gly Asn Val 115 120 125

Thr Ala Thr Arg Asx Ala Ile Thr Thr Thr Pro Thr Thr Thr Ile Thr 130 135 140

Thr Thr Ile Ala Gly Thr Tyr Gln Ser Pro 145 150

- (2) INFORMATION FOR SEQ ID NO:10:
- 30 (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 163 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS:
 - (D) TOPOLOGY: linear
- 35 (ii) MOLECULE TYPE: protein
 - (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Asn Tyr Tyr His Ser Tyr Asn Asn Asn Tyr Tyr Cys Phe His Asp Ile $1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15$

	Thr	Gln	Glu	Arg 20	Gln	Cys	His	Arg	His 25	Glu	Ser	Cys	Tyr	His 30	Tyr	Tyr
	Ala	Asp	Tyr 35	Tyr	Asn	Tyr	Tyr	Asp 40	Tyr	Cys	Arg	Tyr	Val 45	Pro	Ile	Thr
5	Val	Thr 50	Asn	Asn	Thr	Thr	Pro 55	Val	Val	Ser	Glu	Thr 60	Leu	Pro	Ser	Leu
	Pro 65	Val	Lys	Ile	Arg	Asn 70	Lys	Ile	Gly	Ala	Cys 75	Glu	Cys	Gly	Glu	Trp 80
10	Thr	Glu	Trp	Thr	Gly 85	Pro	Cys	Ser	Gln	Glu 90	Cys	Gly	Gly	Cys	Gly 95	Lys
	Arg	Leu	Arg	Thr 100	Arg	Gln	Cys	Ser	Ser 105	Asp	Thr	Glu	Cys	Arg 110	Thr	Glu
	Glu	Lys	Arg 115	Ala	Cys	Ala	Phe	Lys 120	Phe	Ala	His	Thr	Gly 125	Leu	Ile	Ser
15	Leu	Ser 130	Ile	Met	Glu	Ser	Phe 135	Ile	Tyr	Phe	Gly	Arg 140	Ala	Ala	Val	Leu
	Val 145	Tyr	Ser	Asp	Arg	Glu 150	Ile	Cys	Val	Gln	His 155	Leu	Met	Ile	Thr	Arg 160
20	Ile	His	Phe													
	(2) INFO	RMATI	ON E	OR S	SEQ I	D NO	:11	:								
25	(i)	SEQU (A) (B) (C) (D)	LEN TYI STI	NGTH: PE: & RANDE	: 638 amino EDNES	3 ami	ino a id		5							
	(ii)	MOLE	ECULE	E TYP	?E: p	orote	ein									
	(xi)	SEQU	JENCE	E DES	SCRI	OITS	1: SI	EQ II	оио:	:11:						
30	Lys 1	Ala	Ser	Arg	Asn 5	Thr	Leu	Asp	His	Ile 10	Lys	Gln	Leu	Ile	Thr 15	Leu
	Asn	Val	Gln	Arg 20	Glu	Ile	Gly	Asn	Ile 25	Asp	Asp	Lys	Thr	Leu 30	Ala	Asp
	Glu	Ile	Val 35	Leu	Gln	Arg	Arg	Asp 40	Pro	Glu	Ala	Lys	Trp 45	His	His	Asn
35	Glu	Leu 50	Phe	Ile	Asn	Asp	Pro 55	Asp	Ala	Tyr	Tyr	Gln 60	Gly	Asp	Val	Asp

	Leu 65	Ser	Glu	Lys	Gln	Ala 70	Glu	Ile	Leu	Ser	Glu 75	His	Phe	Lys	Leu	Asn 80
	Glu	Ile	Ala	Leu	Thr 85	Glu	Lys	Asp	Asp	Thr 90	Ile	Ile	Arg	Arg	Lys 95	Lys
5	Ser	Ile	Gly	Arg 100	Glu	Pro	Phe	Tyr	Val 105	Arg	Trp	Asn	His	Lys 110	Arg	Pro
	Ile	Ser	Туг 115	Glu	Phe	Ala	Glu	Ser 120	Ile	Pro	Leu	Glu	Thr 125	Arg	Arg	Lys
10	Ile	Arg 130	Ser	Ala	Ile	Ala	Met 135	Trp	Glu	Glu	Arg	Thr 140	Cys	Ile	Arg	Phe
	Gln 145	Glu	Asn	Gly	Pro	Asn 150	Val	Asp	Arg	Ile	Glu 155	Phe	Tyr	Asp	Gly	Gly 160
	Gly	Cys	Ser	Ser	Phe 165	Val	Gly	Arg	Gln	Glu 170	Gly	Ile	Ser	Ile	Ser 175	Thr
15	Pro	Gly	Cys	Asp 180	Ile	Ile	Gly	Ile	Ile 185	Ser	His	Glu	Ile	Gly 190	His	Thr
	Leu	Gly	Ile 195	Phe	His	Glu	Gln	Ala 200	Arg	Arg	Asp	Gln	Lys 205	Asn	His	Ile
20	Phe	Ile 210	Asn	Tyr	Asn	Asn	Ile 215	Pro	Ser	Ser	Arg	Trp 220	Asn	Asn	Phe	Phe
	Pro 225	Leu	Ser	Glu	Tyr	Glu 230	Ala	Asp	Met	Phe	Asn 235	Leu	Pro	Tyr	Asp	Thr 240
	Gly	Ser	Val	Met	His 245	Tyr	Gly	Ser	Tyr	Gly 250	Phe	Ala	Arg	Asn	Pro 255	Tyr
25	Glu	Pro	Thr	Ile 260	Thr	Thr	Arg	Asp	Lys 265	Phe	Gln	Gln	Tyr	Thr 270	Ile	Gly
	Gln	Arg	Glu 275	Gly	Pro	Ser	Phe	Leu 280	Asp	Met	His	Leu	Ile 285	Asn	Ser	Ala
30	Tyr	Arg 290	Cys	Thr	Glu	Gln	Cys 295	Ala	Asp	Met	His	Cys 300	Asp	His	Asn	Gly
	Tyr 305	Pro	Asp	Pro	Asn	Asn 310	Cys	Ala	Lys	Cys	Leu 315	Cys	Pro	Asp	Gly	Phe 320
	Ala	Gly	Arg	Thr	Cys 325	Gln	Phe	Val	Gln	Tyr 330	Thr	Ser	Cys	Gly	Ala 335	Leu
35	Ile	Lys	Ala	Arg 340	Lys	Met	Pro	Val	Thr 345	Ile	Ser	Ser	Pro	Asn 350	Tyr	Pro
	Asn	Phe	Phe 355	Asn	Tyr	Gly	Asp	Gln 360	Cys	Ile	Trp	Leu	Leu 365	Thr	Ala	Pro

		Arg	Val 370	Phe	Val	Asn	Leu	Gln 375	Phe	Val	Glu	Gln	Phe 380	Gln	Leu	Gln	Cys
		Glu 385	Asp	Thr	Cys	Asp	Lys 390	Ser	Tyr	Val	Glu	Val 395	Lys	Ala	Asp	Ala	Asp 400
	5	Phe	Arg	Pro	Thr	Gly 405	Tyr	Arg	Phe	Cys	Cys 410	Ser	Arg	Val	Pro	Arg 415	His
		Ile	Phe	Gln	Ser 420	Ala	Thr	Asn	Glu	Met 425	Val	Val	Il.e	Phe	Arg 430	Gly	Phe
in ente	10	Gly	Asp	Ala 435	Gly	Asn	Gly	Phe	Lys 440	Ala	Lys	Ile	Trp	Ser 445	Asn	Val	Asp
1		Asp	Asp 450	Ile	Ala	Asn	Thr	Ile 455	Val	Thr	Thr	Glu	Met 460	Ala	Lys	Ile	Ser
		Glu 465	Lys	Ile	Pro	Lys	Leu 470	Thr	Val	Pro	Ile	Val 475	Lys	Thr	Ile	Thr	Thr 480
	15	Pro	Thr	Ile	Thr	Thr 485	Thr	Thr	Ala	Phe	Met 490	Ile	Ser	Pro	Lys	Lys 495	Gly
		Asn	Val	Thr	Ala 500	Thr	Arg	Val	Ala	Ile 505	Thr	Thr	Thr	Pro	Thr 510	Thr	Thr
	20	Ile	Thr	Thr 515	Thr	Ile	Ala	Gly	Thr 520	Tyr	Gln	Ser	Val	Thr 525	Asn	Asn	Thr
		Thr	Pro 530	Val	Val	Ser	Glu	Thr 535	Leu	Pro	Ser	Leu	Pro 540	Val	Lys	Ile	Arg
		Asn 545	Lys	Ile	Gly	Ala	Cys 550	Glu	Cys	Gly	Glu	Trp 555	Thr	Glu	Trp	Thr	Gly 560
And	25	Pro	Cys	Ser	Gln	Glu 565	Cys	Gly	Gly	Cys	Gly 570	Lys	Arg	Leu	Arg	Thr 575	Arg
		Gln	Cys	Ser	Ser 580	Asp	Thr	Glu	Cys	Arg 585	Thr	Glu	Glu	Lys	Arg 590	Ala	Суѕ
	30	Ala	Phe	Lys 595	Phe	Ala	His	Thr	Gly 600	Leu	Ile	Ser	Leu	Ser 605	Ile	Met	Glu
		Ser	Phe 610	Ile	Tyr	Phe	Gly	Arg 615	Ala	Ala	Val	Leu	Val 620	Tyr	Ser	Asp	Arg
		Glu 625	Ile	Cys	Val	Gln	His 630	Leu	Met	Ile	Thr	Arg 635	Ile	His	Phe		

	(2) INFORMATION FOR SEQ ID NO:12:	
5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 143 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: DNA (genomic)	
10	(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 1143	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:	
	GGA TCC TGT GGT TCA TGT TGG GCT TTT TCT GTT ACT GGC AAT ATT GCA Gly Ser Cys Gly Ser Cys Trp Ala Phe Ser Val Thr Gly Asn Ile Ala 1 5 10 15	48
15	AGT CTC TGG GCT ATT AAA ACA GGT GAT TTG ATA TCG CTT TCC GAG CAA Ser Leu Trp Ala Ile Lys Thr Gly Asp Leu Ile Ser Leu Ser Glu Gln 20 25 30	96
20	GAA TTG ATT GAT TGT GAT GTG GTT GAT GAG GGC TGC AAC GGC GGC TA Glu Leu Ile Asp Cys Asp Val Val Asp Glu Gly Cys Asn Gly Gly 35 40 45	143
	(2) INFORMATION FOR SEQ ID NO:13:	
25	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 47 amino acids(B) TYPE: amino acid(D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: protein	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:	
	Gly Ser Cys Gly Ser Cys Trp Ala Phe Ser Val Thr Gly Asn Ile Ala 1 5 10 15	
30	Ser Leu Trp Ala Ile Lys Thr Gly Asp Leu Ile Ser Leu Ser Glu Gln 20 25 30	
	Glu Leu Ile Asp Cys Asp Val Val Asp Glu Gly Cys Asn Gly Gly 35 40 45	
	(2) INFORMATION FOR SEQ ID NO:14:	
35	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 18 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	

	(ii) MOLECULE TYPE: DNA (genomic)	
5	<pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION: 12 (D) OTHER INFORMATION: /label= INOSINE</pre>	
	<pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION: 15 (D) OTHER INFORMATION: /label= INOSINE</pre>	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:	
	ACWCATGAAA TNGSNCAT	18
	(2) INFORMATION FOR SEQ ID NO:15:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:	
20	AATACGACTC ACTATAG	17
	(2) INFORMATION FOR SEQ ID NO:16:	
25	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 30 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:	
	TGGTATTATA TCACATGAAA TTGGTCATAC	30
30	(2) INFORMATION FOR SEQ ID NO:17:	
	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 29 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single	
35	(D) TOPOLOGY: linear	
•	(ii) MOLECULE TYPE: DNA (genomic)	

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:	
	CCCAATTGTG TACTGTTGAA ATTTATCAC	29
	(2) INFORMATION FOR SEQ ID NO:18:	
5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 25 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: DNA (genomic)	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:	
	CGGGATCCTG TGGWTCATGY TGGGC	25
	(2) INFORMATION FOR SEQ ID NO:19:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: DNA (genomic)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:	
20	TANCCNCCRT TRCANCCYTC	20
	(2) INFORMATION FOR SEQ ID NO:20:	
25	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 689 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: cDNA	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:	
	TCACATGAAA TTGGTCATAC TTTAGGAATA TTTCATGAGC AAGCACGTCG TGATCAAAAA	60
30	AATCATATTT TTATTAATTA CAACAATATT CCATCAAGCC GTTGGAACAA TTTTTTTCCA	120
	TTATCAGAAT ATGAAGCTGA TATGTTTAAT TTACCTTATG ATACAGGATC AGTAATGCAC	180
	TATGGTTCAT ACGGATTTGC AAGAAATCCG TATGAACCAA CTATTACAAC ACGTGATAAA	240
	TTTCAACAGT ACACAATTGG GCAACGTGAA GGGCCATCAT TTCTGGATTA TGCATCTGTT	300
	AAGCTTTATC TACAAACGCA TTAATGATAT TGTTATCAAA TGGATGATAA TTTCAATAAG	360

10

15

20

25

TATAAACAGC	GCTTATCGTT	GTACAGAACA	ATGTGCTGAT	ATGCACTGCG	ATCATAATGG	420
TTATCCGGAT	CCTAATAATT	GCGCGAAATG	CTTGTGTCCA	GATGGTTTTG	CTGGTCGTAC	480
CTGTCAATTT	GTTCAATATA	CATCTTGCGG	AGCTCTCATT	AAGGTAAGTA	TTGTCTTTTG	540
ACCTCTTCTC	TGACTAAAAT	ATAAGTTAAG	CATATGTATC	TTCCGTCTAA	TGATTTTCTT	600
GATTTTGATT	TGTTCAATGC	TCTTCTTGAT	AATAATATAA	AAATTTTTGA	AAATAAAGTT	660
AACTTTTGGT	САААААААА	АААААААА				689

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 804 base pairs

(B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GATCCTGAGG CAAAATGGCA TCATAATGAA CTATTCATTA ATGATCCAGA TGCATACTAT 60 CAAGGCGATG TCGATTTGTC GGAAAAACAA GCCGAAATTC TAAGCGAACA TTTTAAAAAT 120 GAAATTGCTT TAACAGAGAA AGACGACACA ATAATACGGC GAAAAAAGAG CATTGGTCGT 180 GAACCATTTT ACGTAAGATG GAATCATAAA CGTCCCATTA GCTATGAATT TGCGGAAAGT 240 ATTCCATTAG AAACACGTAG AAAAATTCGT TCAGCAATAG CAATGTGGGA AGAACGAACA 300 TGCATACGAT TCCAAGAAAA TGGCCCAAAT GTAGATCGAA TTGAATTTTA CGACGGTGGC 360 GGTTGTTCAA GTTTTGTCGG CCGAACAGGA GGGAATTTCA ATTTCAACAC CAGGATGTGA 420 TATTATTGGT ATTATATCAC ATGAAATTGG TCATACTTTA GGAATATTTC ATGAGCAAGC 480 ACGTCGTGAT CAAAAAAATC ATATTTTAT TAATTACAAC AATATTCCAT CAAGCCGTTG 540 GAACAATTTT TTTCCATTAT CAGAATATGA AGCTGATATG TTTAATTTAC CTTATGATAC 600 AGGATCAGTA ATGCACTATG GTTCATACGG ATTTGCAAGA AATCCGTATG AACCAACTAT 660 TACAACACGT GATAAATTTC AACAGTACAC AATTGGGCAA CGTGAAGGGC CATCATTTCT 720 GGATTATGCA TCTGATAAAC AGCGCTTATC GTTGTACAGA ACAATGTGCT GATATGCACT 780 GCGATCATAA TGGTTATCCG GATC 804

	5	(A) LENGTH: 271 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
		(ii) MOLECULE TYPE: cDNA	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:	
		TGGTATTATA TCACATGAAA TTGGTCATAC TTTAGGAATA TTTCATGAGC AAGCACGTCG	60
	10	TGATCAAAAA AATCATATTT TTATTAATTA CAACAATATT CCATCAAGCC GTTGGAACAA	120
		TTTTTTCCA TTATCAGAAT ATGAAGCTGA TATGTTTAAT TTACCTTATG ATACAGGATC	180
		AGTAATGCAC TATGGTTCAT ACGGATTTGC AAGAAATCCG TATGAACCAA CTATTACAAC	240
,		ACGTGATAAA TTTCAACAGT ACACAATTGG G	271
		(2) INFORMATION FOR SEQ ID NO:23:	
	15	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 142 base pairs(B) TYPE: nucleic acid(C) STRANDEDNESS: single(D) TOPOLOGY: linear	
- - 	20	(ii) MOLECULE TYPE: cDNA	
		<pre>(ix) FEATURE: (A) NAME/KEY: CDS (B) LOCATION: 3140 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:</pre>	
<u> </u>	25	GA TCC TGT GGT TCA TGT TGG GCT TTT TCT GTT ACT GGC AAT ATT GCA	47
Arri .	25	Ser Cys Gly Ser Cys Trp Ala Phe Ser Val Thr Gly Asn Ile Ala 1 5 10 15	47
	30	AGT CTC TGG GCT ATT AAA ACA GGT GAT TTG ATA TCG CTT TCC GAG CAA Ser Leu Trp Ala Ile Lys Thr Gly Asp Leu Ile Ser Leu Ser Glu Gln 20 25 30	95
		GAA TTG ATT GAT TGT GAT GTG GTT GAT GAG GGC TGC AAC GGC GGC Glu Leu Ile Asp Cys Asp Val Val Asp Glu Gly Cys Asn Gly Gly 35 40 45	140
		TA	142

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

5	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 46 amino acids(B) TYPE: amino acid(D) TOPOLOGY: linear													
	(ii) MOLECULE TYPE: protein													
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:													
	Ser Cys Gly Ser Cys Trp Ala Phe Ser Val Thr Gly Asn Ile Ala Ser 1 5 10 15													
10	Leu Trp Ala Ile Lys Thr Gly Asp Leu Ile Ser Leu Ser Glu Gln Glu 20 25 30													
	Leu Ile Asp Cys Asp Val Val Asp Glu Gly Cys Asn Gly Gly 35 40 45													
	(2) INFORMATION FOR SEQ ID NO:25:													
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 34 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 													
20	<pre>(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "primer"</pre>													
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:													
	GTCGGATCCG CAGGAGGGAA TTTCAATTTC AACA	34												
	(2) INFORMATION FOR SEQ ID NO:26:													
25	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 34 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 													
30	<pre>(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "primer"</pre>													
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:													
	TCAAGATCTA ATCCAGAAAT GATGGCCCTT CACG	34												

(2) INFORMATION FOR SEQ ID NO:24:

	(2) INFORMATION FOR SEQ ID NO:27:	
5	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 19 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	<pre>(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "primer"</pre>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:	
10	GGAAACAGCT ATGACCATG	19
	(2) INFORMATION FOR SEQ ID NO:28:	
15	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 17 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	<pre>(ii) MOLECULE TYPE: other nucleic acid (A) DESCRIPTION: /desc = "primer"</pre>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:	
20	GTAAAACGAC GGCCAGT	17
	(2) INFORMATION FOR SEQ ID NO:29:	
25	 (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 2292 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear 	
	(ii) MOLECULE TYPE: cDNA	
30	<pre>(ix) FEATURE: (A) NAME/KEY: misc_feature (B) LOCATION: 1459 (D) OTHER INFORMATION: /note= "1459:S=C or G; aa463:Xaa=Alanine or Glycine"</pre>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:	
	TAGATTTCGA TTCGTCTTTG TTAATTCATC TTCGTCAGAT TTATTAGAGA AAAATAAAAA	60
35	TTTTGATCGC AATGAAGCAG GTTATCATCT TTCCTCAGCT TTTCATTTGT TTCATTGTTC	120
	AGTCAGTGGA AAATTATCGA ACGCAGAAAG CATCACGAAA TACGTTAGAT CACATCAAAC	180
	AACTTATCAC CTTGAACGTA CAAAGAGAGA TTGGAAACAT AGATGATAAG ACATTAGCTG	240

						•
300	GAACTATTCA	GCATCATAAT	AGGCAAAATG	CGGGATCCTG	ATTACAACGA	ATGAAATAGT
360	CAAGCCGAAA	GTCGGAAAAA	ATGTCGATTT	TATCAAGGCG	AGATGCATAC	TTAATGATCC
420	ACAATAATAC	GAAAGACGAC	CTTTAACAGA	AATGAAATTG	ACATTTTAAA	TTCTAAGCGA
480	AAACGTCCCA	ATGGAATCAT	TTTACGTAAG	CGTGAACCAT	GAGCATTGGT	GGCGAAAAA
540	CGTTCAGCAA	TAGAAAAATT	TAGAAACACG	AGTATTCCAT	ATTTGCGGAA	TTAGCTATGA
600	AATGTTGATC	AAATGGCCCA	GATTCCAAGA	ACATGCATAC	GGAAGAACGA	TAGCAATGTG
660	GGAGGCACGC	CGGCCGAACA	CAAGTTTTGT	GGCGGTTGTT	TTACGACGGT	GAATTGAATT
720	CATGAAATTG	TATTATATCA	ATATTATTGG	CCAGGATGTG	AATTTCAACA	AAGGAATTTC
780	CATATTTTTA	TCAAAAAAAT	CACGTCGTGA	CATGAGCAAG	AGGAATATTT	GTCATACTTT
840	TCAGAATATG	TTTTCCATTA	GGAACAATTT	TCAAGCCGTT	CAATATTCCA	TTAATTACAA
900	GGTTCATACG	AATGCACTAT	CAGGATCAGT	CCTTATGATA	GTTTAATTTA	AAGCTGATAT
960	CAACAGTACA	TGATAAATTT	TTACAACACG	GAACCAACTA	AAATCCGTAT	GATTTGCAAG
1020	AGCGCTTATC	ATCTATAAAC	TGGATTATGC	CCATCATTTC	ACGTGAAGGG	CAATTGGGCA
1080	GATCCTAATA	TGGTTATCCG	GCGATCATAA	GATATGCACT	ACAATGTGCT	GTTGTACAGA
1140	TTTGTTCAAT	TACCTGTCAA	TTGCTGGTCG	CCAGATGGTT	ATGCTTGTGT	ATTGCGCGAA
1200	TCGAGCCCAA	TGTTACGATT	GGAAAATGCC	ATTAAGGCGA	CGGAGCTCTC	ATACATCTTG
12.60	GCTCCACGCG	GTTGCTTACA	AATGTATTTG	GTTGGTGATC	CTTCTTCAAT	ATTATCCAAA
1320	GATACGTGTG	ACAATGTGAA	AATTTCAATT	TTTGTTGAAC	AAATTTGCAG	GTGGATTCGT
1380	TATCGATTTT	ACCTACTGGA	CTGATTTTCG	AAAGCTGACG	TGTAGAAGTG	ATAAATCCTA
1440	GTAGTAATAT	AAACGAGATG	AATCTGCGAC	CATATTTTTC	AGTGCCACGT	GTTGTTCGCG
1500	AACGTAGATG	AATTTGGTCA	TTAAAGCTAA	GGAAATGGCT	TGGTGATGCG	TTCGCGGTTT
1560	AAAATACCGA	AATTTCGGAA	AAATGGCAAA	GTAACAACTG	TAATACAATT	ATGATATAGC
1620	ACTACTGCTT	AATAACAACT	CCACTCCTAC	AAAACTATTA	TCCAATAGTT	AGCTAACAGT
1680	ACTACTACGC	AGTTGCTATC	CCGCCACGAG	GGCAATGTCA	ACCCAAGAAA	TCATGATATC
1740	ACTAATAATA	AATCACCGTA	GTACGGTACC	ACTATTGCCG	AATTACTACG	CGACTACTAC
1800	AACAAAATAG	CAAGATTCGA	CATTGCCAGT	ACTTTACCAT	AGTAAGTGAA	CTACACCTGT
1860	GAATGTGGCG	ATGCTCTCAA	GGACAGGTCC	TGGACAGAAT	ATGTGGTGAA	GTGCATGCGA
1920	AGAACAGAAG	TACGGAATGT	GTTCATCAGA	ACACGTCAGT	ACGTCTTCGA	GTTGCGGAAA

AAAA	ACGTGC GTGT	GCTTTT AAAG	TTGCC CAT	ACGGGAC	TAATTTCCTT	ATCAATAATG	1980
GAGA	GTTTCA TATA	CTTTGG AAGG	SCTGCT GTO	STTGGTCT	ATTCCGATCG	GGAGATATGT	2040
GTTC	AGCACT TGAT	GATAAC GAGA	ATCCAT TTO	TGAAATT	TCTAGAATCA	CTGTTGAACA	2100
TGCA	AGATTC TCGA	AAAAAC GATA	ATTTGC CT	ACTCGAA	AAAGAAGTGA	TTGAATGATT	2160
CGAT	AATATT GATT	AATAAA ACGG	STTGTA TTO	CTCGTCAT	AGAGTATCCG	TTGATGTTTT	2220
TATC	CAAAAA ATTC	TCTTGC TTTT	AATTAT TGI	GAATAAA	ACTTTTGTTT	АСССАААААА	2280
AAAA	AA AAAAA						2292
(2)	INFORMATION	N FOR SEQ I	NO:30:				
	(A) I (B) I (C) S	NCE CHARACTI LENGTH: 207 TYPE: nucle: STRANDEDNES: TOPOLOGY: 1:	5 base pai ic acid 3: single	rs			
	(ii) MOLECU	JLE TYPE: cl	ANC				
		RE: NAME/KEY: C LOCATION: 1					
	(xi) SEQUEN	NCE DESCRIP	rion: SEQ	ID NO:30	:		
		ATC ATC TT L lle lle Pl 5					48
		AAT TAT CG 1 Asn Tyr A)		Lys Ala	Ser Arg As		96
		CAA CTT ATO Gln Leu I					144
		AAG ACA TTA Lys Thr Lo					192
		AAA TGG CA A Lys Trp H: 70			Phe Ile As		240
		CAA GGC GA' c Gln Gly A: 85					288

	ATT CTA AGC GAA CAT TTT AAA AAT GAA ATT GCT TTA ACA GAG AAA GAC Ile Leu Ser Glu His Phe Lys Asn Glu Ile Ala Leu Thr Glu Lys Asp 100 105 110	336
5	GAC ACA ATA ATA CGG CGA AAA AAG AGC ATT GGT CGT GAA CCA TTT TAC Asp Thr Ile Ile Arg Arg Lys Lys Ser Ile Gly Arg Glu Pro Phe Tyr 115 120 125	384
	GTA AGA TGG AAT CAT AAA CGT CCC ATT AGC TAT GAA TTT GCG GAA AGT Val Arg Trp Asn His Lys Arg Pro Ile Ser Tyr Glu Phe Ala Glu Ser 130 135 140	432
10	ATT CCA TTA GAA ACA CGT AGA AAA ATT CGT TCA GCA ATA GCA ATG TGG Ile Pro Leu Glu Thr Arg Arg Lys Ile Arg Ser Ala Ile Ala Met Trp 145 150 155 160	480
15	GAA GAA CGA ACA TGC ATA CGA TTC CAA GAA AAT GGC CCA AAT GTT GAT Glu Glu Arg Thr Cys Ile Arg Phe Gln Glu Asn Gly Pro Asn Val Asp 165 170 175	528
	CGA ATT GAA TTT TAC GAC GGT GGC GGT TGT TCA AGT TTT GTC GGC CGA Arg Ile Glu Phe Tyr Asp Gly Gly Cys Ser Ser Phe Val Gly Arg 180 185 190	576
20	ACA GGA GGC ACG CAA GGA ATT TCA ATT TCA ACA CCA GGA TGT GAT ATT Thr Gly Gly Thr Gln Gly Ile Ser Ile Ser Thr Pro Gly Cys Asp Ile 195 200 205	624
	ATT GGT ATT ATA TCA CAT GAA ATT GGT CAT ACT TTA GGA ATA TTT CAT Ile Gly Ile Ile Ser His Glu Ile Gly His Thr Leu Gly Ile Phe His 210 220	672
25	GAG CAA GCA CGT CGT GAT CAA AAA AAT CAT ATT TTT ATT AAT TAC AAC Glu Gln Ala Arg Arg Asp Gln Lys Asn His Ile Phe Ile Asn Tyr Asn 225 230 235 240	720
30	AAT ATT CCA TCA AGC CGT TGG AAC AAT TTT TTT CCA TTA TCA GAA TAT Asn Ile Pro Ser Ser Arg Trp Asn Asn Phe Phe Pro Leu Ser Glu Tyr 245 250 255	768
	GAA GCT GAT ATG TTT AAT TTA CCT TAT GAT ACA GGA TCA GTA ATG CAC Glu Ala Asp Met Phe Asn Leu Pro Tyr Asp Thr Gly Ser Val Met His 260 265 270	816
35	TAT GGT TCA TAC GGA TTT GCA AGA AAT CCG TAT GAA CCA ACT ATT ACA Tyr Gly Ser Tyr Gly Phe Ala Arg Asn Pro Tyr Glu Pro Thr Ile Thr 275 280 285	864
	ACA CGT GAT AAA TTT CAA CAG TAC ACA ATT GGG CAA CGT GAA GGG CCA Thr Arg Asp Lys Phe Gln Gln Tyr Thr Ile Gly Gln Arg Glu Gly Pro 290 295 300	912
40	TCA TTT CTG GAT TAT GCA TCT ATA AAC AGC GCT TAT CGT TGT ACA GAA Ser Phe Leu Asp Tyr Ala Ser Ile Asn Ser Ala Tyr Arg Cys Thr Glu 305 310 315 320	960

	CAA TGT GCT GAT ATG CAC TGC GAT CAT AAT GGT TAT CCG GAT CCT AAT Gln Cys Ala Asp Met His Cys Asp His Asn Gly Tyr Pro Asp Pro Asn 325 330 335	1008
5	AAT TGC GCG AAA TGC TTG TGT CCA GAT GGT TTT GCT GGT CGT ACC TGT Asn Cys Ala Lys Cys Leu Cys Pro Asp Gly Phe Ala Gly Arg Thr Cys 340 345 350	1056
	CAA TTT GTT CAA TAT ACA TCT TGC GGA GCT CTC ATT AAG GCG AGG AAA Gln Phe Val Gln Tyr Thr Ser Cys Gly Ala Leu Ile Lys Ala Arg Lys 355 360 365	1104
10	ATG CCT GTT ACG ATT TCG AGC CCA AAT TAT CCA AAC TTC TTC AAT GTT Met Pro Val Thr Ile Ser Ser Pro Asn Tyr Pro Asn Phe Phe Asn Val 370 380	1152
15	GGT GAT CAA TGT ATT TGG TTG CTT ACA GCT CCA CGC GGT GGA TTC GTA Gly Asp Gln Cys Ile Trp Leu Leu Thr Ala Pro Arg Gly Gly Phe Val 385 390 395 400	1200
	AAT TTG CAG TTT GTT GAA CAA TTT CAA TTA CAA TGT GAA GAT ACG TGT Asn Leu Gln Phe Val Glu Gln Phe Gln Leu Gln Cys Glu Asp Thr Cys 405 410 415	1248
20	GAT AAA TCC TAT GTA GAA GTG AAA GCT GAC GCT GAT TTT CGA CCT ACT Asp Lys Ser Tyr Val Glu Val Lys Ala Asp Ala Asp Phe Arg Pro Thr 420 425 430	1296
	GGA TAT CGA TTT TGT TGT TCG CGA GTG CCA CGT CAT ATT TTT CAA TCT Gly Tyr Arg Phe Cys Cys Ser Arg Val Pro Arg His Ile Phe Gln Ser 435 440 445	1344
25	GCG ACA AAC GAG ATG GTA GTA ATA TTT CGC GGT TTT GGT GAT GCG GGA Ala Thr Asn Glu Met Val Val Ile Phe Arg Gly Phe Gly Asp Ala Gly 450 455 460	1392
30	AAT GGC TTT AAA GCT AAA ATT TGG TCA AAC GTA GAT GAT GAT ATA GCT Asn Gly Phe Lys Ala Lys Ile Trp Ser Asn Val Asp Asp Asp Ile Ala 465 470 475. 480	1440
	AAT ACA ATT GTA ACA ACT GAA ATG GCA AAA ATT TCG GAA AAA ATA CCG Asn Thr Ile Val Thr Thr Glu Met Ala Lys Ile Ser Glu Lys Ile Pro 485 490 495	1488
35	AAG CTA ACA GTT CCA ATA GTT AAA ACT ATT ACC ACT CCT ACA ATA ACA Lys Leu Thr Val Pro Ile Val Lys Thr Ile Thr Thr Pro Thr Ile Thr 500 505 510	1536
	ACT ACT ACT GCT TTC ATG ATA TCA CCC AAG AAA GGC AAT GTC ACC GCC Thr Thr Ala Phe Met Ile Ser Pro Lys Lys Gly Asn Val Thr Ala 515 520 525	1584
40	ACG AGA GTT GCT ATC ACT ACT ACG CCG ACT ACT ACA ATT ACT ACG ACT Thr Arg Val Ala Ile Thr Thr Thr Pro Thr Thr Thr Ile Thr Thr Thr 530 540	1632

		GT ACG GTA CCA ATC ACC GTA ACT AAT AAT ACT ACA CCT GTA Lly Thr Val Pro Ile Thr Val Thr Asn Asn Thr Thr Pro Val 550 555 560	1680
5		AA ACT TTA CCA TCA TTG CCA GTC AAG ATT CGA AAC AAA ATA Flu Thr Leu Pro Ser Leu Pro Val Lys Ile Arg Asn Lys Ile 565 570 575	1728
		GC GAA TGT GGT GAA TGG ACA GAA TGG ACA GGT CCA TGC TCT Tys Glu Cys Gly Glu Trp Thr Glu Trp Thr Gly Pro Cys Ser 580 590	1776
10	Gln Glu C	GT GGC GGT TGC GGA AAA CGT CTT CGA ACA CGT CAG TGT TCA Tys Gly Gly Cys Gly Lys Arg Leu Arg Thr Arg Gln Cys Ser 95 600 605	1824
15		CG GAA TGT AGA ACA GAA GAA AAA CGT GCG TGT GCT TTT AAA hr Glu Cys Arg Thr Glu Glu Lys Arg Ala Cys Ala Phe Lys 615 620	1872
		CA TAC GGG ACT AAT TTC CTT ATC AAT AAT GGA GAG TTT CAT TO Tyr Gly Thr Asn Phe Leu Ile Asn Asn Gly Glu Phe His 630 635 640	1920
20		GG AAG GGC TGC TGT GTT GGT CTA TTC CGA TCG GGA GAT ATG rp Lys Gly Cys Cys Val Gly Leu Phe Arg Ser Gly Asp Met 645 650 655	1968
		CA CTT GAT GAT AAC GAG AAT CCA TTT CTG AAA TTT CTA GAA la Leu Asp Asp Asn Glu Asn Pro Phe Leu Lys Phe Leu Glu 660 665 670	2016
25	Ser Leu L	TG AAC ATG CAA GAT TCT CGA AAA AAC GAT AAT TTG CCT GAC eu Asn Met Gln Asp Ser Arg Lys Asn Asp Asn Leu Pro Asp 685	2064
30	TCG AAA AA Ser Lys L 690		2076

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 692 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Met Lys Gln Val Ile Ile Phe Pro Gln Leu Phe Ile Cys Phe Ile Val 1 5 10 15

	Gln	Ser	Val	Glu 20	Asn	Tyr	Arg	Thr	Gln 25	Lys	Ala	Ser	Arg	Asn 30	Thr	Leu
	Asp	His	Ile 35	Lys	Gln	Leu	Ile	Thr 40	Leu	Asn	Val	Gln	Arg 45	Glu	Ile	Gly
5	Asn	Ile 50	Asp	Asp	Lys	Thr	Leu 55	Ala	Asp	Glu	Ile	Val 60	Leu	Gln	Arg	Arg
	Asp 65	Pro	Glu	Ala	Lys	Trp 70	His	His	Asn	Glu	Leu 75	Phe	Ile	Asn	Asp	Pro 80
10	Asp	Ala	Tyr	Tyr	Gln 85	Gly	Asp	Val	Asp	Leu 90	Ser	Glu	Lys	Gln	Ala 95	Glu
	Ile	Leu	Ser	Glu 100	His	Phe	Lys	Asn	Glu 105	Ile	Ala	Leu	Thr	Glu 110	Lys	Asp
	Asp	Thr	Ile 115	Ile	Arg	Arg	Lys	Lys 120	Ser	Ile	Gly	Arg	Glu 125	Pro	Phe	Tyr
15	Val	Arg 130	Trp	Asn	His	Lys	Arg 135	Pro	Ile	Ser	Tyr	Glu 140	Phe	Ala	Glu	Ser
	Ile 145	Pro	Leu	Glu	Thr	Arg 150	Arg	Lys	Ile	Arg	Ser 155	Ala	Ile	Ala	Met	Trp 160
20	Glu	Glu	Arg	Thr	Cys 165	Ile	Arg	Phe	Gln	Glu 170	Asn	Gly	Pro	Asn	Val 175	Asp
	Arg	Ile	Glu	Phe 180	Tyr	Asp	Gly	Gly	Gly 185	Cys	Ser	Ser	Phe	Val 190	Gly	Arg
	Thr	Gly	Gly 195	Thr	Gln	Gly	Ile	Ser 200	Ile	Ser	Thr	Pro	Gly 205	Cys	Asp	Ile
25	Ile	Gly 210	Ile	Ile	Ser	His	Glu 215	Ile	Gly	His	Thr	Leu 220	Gly	Ile	Phe	His
	Glu 225	Gln	Ala	Arg	Arg	Asp 230	Gln	Lys	Asn	His	Ile 235	Phe	Ile	Asn	Tyr	Asn 240
30	Asn	Ile	Pro	Ser	Ser 245	Arg	Trp	Asn	Asn	Phe 250	Phe	Pro	Leu	Ser	Glu 255	Tyr
	Glu	Ala	Asp	Met 260	Phe	Asn	Leu	Pro	Tyr 265	Asp	Thr	Gly	Ser	Val 270	Met	His
	Tyr	Gly	Ser 275	Tyr	Gly	Phe	Ala	Arg 280	Asn	Pro	Tyr	Glu	Pro 285	Thr	Ile	Thr
35	Thr	Arg 290	Asp	Lys	Phe	Gln	Gln 295	Tyr	Thr	Ile	Gly	Gln 300	Arg	Glu	Gly	Pro

			•													
	Ser 305	Phe	Leu	Asp	Tyr	Ala 310	Ser	Ile	Asn	Ser	Ala 315	Tyr	Arg	Cys	Thr	Glu 320
	Gln	Суѕ	Ala	Asp	Met 325	His	Cys	Asp	His	Asn 330	Gly	Tyr	Pro	Asp	Pro 335	Asn
5	Asn	Cys	Ala	Lys 340	Cys	Leu	Cys	Pro	Asp 345	Gly	Phe	Ala	Gly	Arg 350	Thr	Cys
	Gln	Phe	Val 355	Gln	Tyr	Thr	Ser	Cys 360	Gly	Ala	Leu	Ile	Lys 365	Ala	Arg	Lys
10	Met	Pro 370	Val	Thr	Ile	Ser	Ser 375	Pro	Asn	Tyr	Pro	Asn 380	Phe	Phe	Asn	Val
	Gly 385	Asp	Gln	Cys	Ile	Trp 390	Leu	Leu	Thr	Ala	Pro 395	Arg	Gly	Gly	Phe	Val 400
	Asn	Leu	Gln	Phe	Val 405	Glu	Gln	Phe	Gln	Leu 410	Gln	Cys	Glu	Asp	Thr 415	Cys
15	Asp	Lys	Ser	Tyr 420	Val	Glu	Val	Lys	Ala 425	Asp	Ala	Asp	Phe	Arg 430	Pro	Thr
	Gly	Tyr	Arg 435	Phe	Cys	Cys	Ser	Arg 440	Val	Pro	Arg	His	Ile 445	Phe	Gln	Ser
20	Ala	Thr 450	Asn	Glu	Met	Val	Val 455	Ile	Phe	Arg	Gly	Phe 460	Gly	Asp	Ala	Gly
	Asn 465	Gly	Phe	Lys	Ala	Lys 470	Ile	Trp	Ser	Asn	Val 475	Asp	Asp	Asp	Ile	Ala 480
	Asn	Thr	Ile	Val	Thr 485	Thr	Glu	Met	Ala	Lys 490	Ile	Ser	Glu	Lys	Ile 495	Pro
25	Lys	Leu	Thr	Val 500	Pro	Ile	Val	Lys	Thr 505	Ile	Thr	Thr	Pro	Thr 510	Ile	Thr
	Thr	Thr	Thr 515	Ala	Phe	Met	Ile	Ser 520	Pro	Lys	Lys	Gly	Asn 525	Val	Thr	Ala
30	Thr	Arg 530	Val	Ala	Ile	Thr	Thr 535	Thr	Pro	Thr	Thr	Thr 540	Ile	Thr	Thr	Thr
	Ile 545	Ala	Gly	Thr	Val	Pro 550	Ile	Thr	Val	Thr	Asn 555	Asn	Thr	Thr	Pro	Val 560
	Val	Ser	Glu	Thr	Leu 565	Pro	Ser	Leu	Pro	Val 570	Lys	Ile	Arg	Asn	Lys 575	Ile
35	Gly	Ala	Cys	Glu 580	Cys	Gly	Glu	Trp	Thr 585	Glu	Trp	Thr	Gly	Pro 590	Cys	Ser

	Gln	Glu	Cys 595	Gly	Gly	Cys	Gly	Lys 600	Arg	Leu	Arg	Thr	Arg 605	Gln	Cys	Ser		
	Ser	Asp 610	Thr	Glu	Cys	Arg	Thr 615	Glu	Glu	Lys	Arg	Ala 620	Cys	Ala	Phe	Lys		
5	Val 625	Cys	Pro	Tyr	Gly	Thr 630	Asn	Phe	Leu	Ile	Asn 635	Asn	Gly	Glu	Phe	His 640		
	Ile	Leu	Trp	Lys	Gly 645	Cys	Cys	Val	Gly	Leu 650	Phe	Arg	Ser	Gly	Asp 655	Met		
10	Cys	Ser	Ala	Leu 660	Asp	Asp	Asn	Glu	Asn 665	Pro	Phe	Leu	Lys	Phe 670	Leu	Glu		
	Ser	Leu	Leu 675	Asn	Met	Gln	Asp	Ser 680	Arg	Lys	Asn	Asp	Asn 685	Leu	Pro	Asp		
	Ser	Lys 690	Lys	Lys														
15	(2)	INFO	RMAT	rion	FOR	SEQ	ID N	10:32	2:									
		(i)	(<i>I</i>	A) LI 3) Ti	CE CH ENGTH (PE:	1: 20 nucl	32 k Leic	ase acio	pai:	îs								
20					rani Polo				gle									
		(ii)	мої	LECUI	LE TY	PE:	cDNA	Ą										
		(xi)	SEÇ	QUENC	CE DE	ESCRI	PTIC	ON: S	SEQ I	D NO	0:32	:						
	TCAG	TCAG	TG G	AAAA	TTAT	C GAZ	ACGCI	AGAA	AGC	ATCAC	CGA A	ATA	CGTTA	AG AT	CAC	ATCAA	60	
	ACAA	ACTTA	TC A	.CCTT	GAAC	G TAC	CAAAC	SAGA	GATI	'GGAI	AAC A	ATAGA	ATGAT	A A	SACAT	TAGC	120	
25	TGAT	'GAAA	TA G	TATT.	ACAA	C GAG	CGGGI	ATCC	TGAG	GCA	AAA :	rggcz	ATCAT	ra at	rgaac	СТАТТ	180	
	CATI	AATG	AT C	CAGA	TGCA:	r act	TATC	AAGG	CGAT	GTC	AT S	rtgt	CGGAI	AA AA	ACAA	GCCGA	240	
	TTAA	CTAA	.GC G	AACA	TTTTI	AAA A	AATGI	TAA	TGCT	ATT	ACA (GAGA	AAGAG	CG AC	CACA	TAAT	300	
	ACGG	CGAA	AA A	AGAG	CATT	G GTO	CGTGF	AACC	ATTI	TAC	TA A	AGAT	GAA7	C A	PAAA	CGTCC	360	
	CATI	AGCT	AT G	AATT	TGCG	G AAA	\GTA1	TTCC	ATTA	AGAAA	ACA (CGTA	SAAAA	AA TI	CGT	CAGC	420	
30	AATA	GCAA	TG T	GGGA	AGAA	C GAA	CATO	CAT	ACG	ATTCO	CAA (SAAA	ATGG	CC CA	AAAC	TAGA	480	
	TCGA	ATTG	та т	TTAA	CGAC	G GTC	GCGC	FT T G	TTCF	AGTI	TTT (FTCG	GCCGI	AA CA	AGGA	GGCAC	540	
	GCCA	GGAA	тт т	CAAT	TTCA	A CAG	CCAG	ATG	TGAT	TTTA	ATT (GTA:	TAT	AT CA	ACATO	SAAAT	600	
	TGGT	CATA	CT T	TAGG	AATA:	r TTC	CATGA	AGCA	AGC	ACGTO	CGT (SATC	AAAA	LA A	CAT	TTTT	660	
	TATI	TTAAT	AC A	ACAA'	TATT	CAT	CAA	SCCG	TTG	AAC	AAT :	rttt	TCC	AT TA	ATCA	GAATA	720	

10

15

20

25

TGAAGCTGAT	ATGTTTAATT	TACCTTATGA	TACAGGATCA	GTAATGCACT	ATGGTTCATA	780
CGGATTTGCA	AGAAATCCGT	ATGAACCAAC	TATTACAACA	CGTGATAAAT	TTCAACAGTA	840
CACAATTGGG	CAACGTGAAG	GGCCATCATT	TCTGGATTAT	GCATCTATAA	ACAGCGCTTA	900
TCGTTGTACA	GAACAATGTG	CTGATATGCA	CTGCGATCAT	AATGGTTATC	CGGATCCTAA	960
TAATTGCGCG	AAATGCTTGT	GTCCAGATGG	TTTTGCTGGT	CGTACCTGTC	AATTTGTTCA	1020
ATATACATCT	TGCGGAGCTC	TCATTAAGGC	GAGGAAAATG	CCTGTTACGA	TTTCGAGCCC	1080
AAATTATCCA	AACTTCTTCA	ATGTTGGTGA	TCAATGTATT	TGGTTGCTTA	CAGCTCCACG	1140
CGGTGGATTC	GTAAATTTGC	AGTTTGTTGA	ACAATTTCAA	TTACAATGTG	AAGATACGTG	1200
TGATAAATCC	TATGTAGAAG	TGAAAGCTGA	CGCTGATTTT	CGACCTACTG	GATATCGATT	1260
TTGTTGTTCG	CGAGTGCCAC	GTCATATTTT	TCAATCTGCG	ACAAACGAGA	TGGTAGTAAT	1320
ATTTCGCGGT	TTTGGTGGTG	CGGGAAATGG	CTTTAAAGCT	AAAATTTGGT	CAAACGTAGA	1380
TGATGATATA	GCTAATACAA	TTGTAACAAC	TGAAATGGCA	AAAATTTCGG	AAAAAATACC	1440
GAAGCTAACA	GTTCCAATAG	TTAAAACTAT	TACCACTCCT	ACAATAACAA	CTACTACTGC	1500
TTTCATGATA	TCACCCAAGA	AAGGCAATGT	CACCGCCACG	AGAGTTGCTA	TCACTACTAC	1560
GCCGACTACT	ACAATTACTA	CGACTATTGC	CGGTACGGTA	CCAATCACCG	ТААСТААТАА	1620
TACTACCCCT	GTAGTAAGTG	AAACTTTACC	ATCATTGCCA	GTCAAGATTC	GAAACAAAAT	1680
AGGTGCATGC	GAATGTGGTG	AATGGACAGA	ATGGACAGGT	CCATGCTCTC	AAGAATGTGG	1740
CGGTTGCGGA	AAACGTCTTC	GAACACGTCA	GTGTTCATCA	GATACGGAAT	GTAGAACAGA	1800
AGAAAAACGT	GCGTGTGCTT	TTAAAGTTTG	CCCATACGGG	ACTAATTTCC	TTATCAATAA	1860
TGGAGAGTTT	CATATACTTT	GGAAGGGCTG	CTGTGTTGGT	CTATTCCGAT	CGGGAGATAT	1920
GTGTTCAGCA	CTTGATGATA	ACGAGAATCC	ATTTCTGAAA	TTTCTAGAAT	CACTGTTGAA	1980
CATGCAAGAT	TCTCGAAAAA	ACGATAATTT	GCCTGACTCG	AAAAAGAAGT	GA	2032

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2028 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..2028

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

5	CAG TCA GTG GAA AAT TAT CGA ACG CAG AAA GCA TCA CGA AAT ACG TTA Gln Ser Val Glu Asn Tyr Arg Thr Gln Lys Ala Ser Arg Asn Thr Leu 1 5 10	48
10	GAT CAC ATC AAA CAA CTT ATC ACC TTG AAC GTA CAA AGA GAG ATT GGA Asp His Ile Lys Gln Leu Ile Thr Leu Asn Val Gln Arg Glu Ile Gly 20 25 30	96
	AAC ATA GAT GAT AAG ACA TTA GCT GAT GAA ATA GTA TTA CAA CGA CGG Asn Ile Asp Asp Lys Thr Leu Ala Asp Glu Ile Val Leu Gln Arg Arg 35 40 45	144
15	GAT CCT GAG GCA AAA TGG CAT CAT AAT GAA CTA TTC ATT AAT GAT CCA Asp Pro Glu Ala Lys Trp His His Asn Glu Leu Phe Ile Asn Asp Pro 50 55 60	192
	GAT GCA TAC TAT CAA GGC GAT GTC GAT TTG TCG GAA AAA CAA GCC GAA Asp Ala Tyr Tyr Gln Gly Asp Val Asp Leu Ser Glu Lys Gln Ala Glu 65 70 75 80	240
20	ATT CTA AGC GAA CAT TTT AAA AAT GAA ATT GCT TTA ACA GAG AAA GAC Ile Leu Ser Glu His Phe Lys Asn Glu Ile Ala Leu Thr Glu Lys Asp 85 90 95	288
25	GAC ACA ATA ATA CGG CGA AAA AAG AGC ATT GGT CGT GAA CCA TTT TAC Asp Thr Ile Ile Arg Arg Lys Lys Ser Ile Gly Arg Glu Pro Phe Tyr 100 105 110	336
	GTA AGA TGG AAT CAT AAA CGT CCC ATT AGC TAT GAA TTT GCG GAA AGT Val Arg Trp Asn His Lys Arg Pro Ile Ser Tyr Glu Phe Ala Glu Ser 115 120 125	384
30	ATT CCA TTA GAA ACA CGT AGA AAA ATT CGT TCA GCA ATA GCA ATG TGG Ile Pro Leu Glu Thr Arg Arg Lys Ile Arg Ser Ala Ile Ala Met Trp 130 135 140	432
	GAA GAA CGA ACA TGC ATA CGA TTC CAA GAA AAT GGC CCA AAC GTA GAT Glu Glu Arg Thr Cys Ile Arg Phe Gln Glu Asn Gly Pro Asn Val Asp 145 150 155 160	480
35	CGA ATT GTA TTT AAC GAC GGT GGC GGT TGT TCA AGT TTT GTC GGC CGA Arg Ile Val Phe Asn Asp Gly Gly Gly Cys Ser Ser Phe Val Gly Arg 165 170 175	528
40	ACA GGA GGC ACG CCA GGA ATT TCA ATT TCA ACA CCA GGA TGT GAT ATT Thr Gly Gly Thr Pro Gly Ile Ser Ile Ser Thr Pro Gly Cys Asp Ile 180 185	576

	•	
	ATT GGT ATT ATA TCA CAT GAA ATT GGT CAT ACT TTA GGA ATA TTT CAT Ile Gly Ile Ile Ser His Glu Ile Gly His Thr Leu Gly Ile Phe His 195 200 205	624
5	GAG CAA GCA CGT CGT GAT CAA AAA AAT CAT ATT TTT ATT AAT TAC AAC Glu Gln Ala Arg Arg Asp Gln Lys Asn His Ile Phe Ile Asn Tyr Asn 210 215 220	672
	AAT ATT CCA TCA AGC CGT TGG AAC AAT TTT TTT CCA TTA TCA GAA TAT Asn Ile Pro Ser Ser Arg Trp Asn Asn Phe Phe Pro Leu Ser Glu Tyr 225 230 235 240	720
10	GAA GCT GAT ATG TTT AAT TTA CCT TAT GAT ACA GGA TCA GTA ATG CAC Glu Ala Asp Met Phe Asn Leu Pro Tyr Asp Thr Gly Ser Val Met His 245 250 255	768
15	TAT GGT TCA TAC GGA TTT GCA AGA AAT CCG TAT GAA CCA ACT ATT ACA Tyr Gly Ser Tyr Gly Phe Ala Arg Asn Pro Tyr Glu Pro Thr Ile Thr 260 270	816
	ACA CGT GAT AAA TTT CAA CAG TAC ACA ATT GGG CAA CGT GAA GGG CCA Thr Arg Asp Lys Phe Gln Gln Tyr Thr Ile Gly Gln Arg Glu Gly Pro 275 280 285	864
20	TCA TTT CTG GAT TAT GCA TCT ATA AAC AGC GCT TAT CGT TGT ACA GAA Ser Phe Leu Asp Tyr Ala Ser Ile Asn Ser Ala Tyr Arg Cys Thr Glu 290 295 300	912
	CAA TGT GCT GAT ATG CAC TGC GAT CAT AAT GGT TAT CCG GAT CCT AAT Gln Cys Ala Asp Met His Cys Asp His Asn Gly Tyr Pro Asp Pro Asn 305 310 315 320	960
25	AAT TGC GCG AAA TGC TTG TGT CCA GAT GGT TTT GCT GGT CGT ACC TGT Asn Cys Ala Lys Cys Leu Cys Pro Asp Gly Phe Ala Gly Arg Thr Cys 325 330 335	1008
30	CAA TTT GTT CAA TAT ACA TCT TGC GGA GCT CTC ATT AAG GCG AGG AAA Gln Phe Val Gln Tyr Thr Ser Cys Gly Ala Leu Ile Lys Ala Arg Lys 340 345 350	1056
	ATG CCT GTT ACG ATT TCG AGC CCA AAT TAT CCA AAC TTC TTC AAT GTT Met Pro Val Thr Ile Ser Ser Pro Asn Tyr Pro Asn Phe Phe Asn Val 355 360 365	1104
35	GGT GAT CAA TGT ATT TGG TTG CTT ACA GCT CCA CGC GGT GGA TTC GTA Gly Asp Gln Cys Ile Trp Leu Leu Thr Ala Pro Arg Gly Gly Phe Val 370 375 380	1152
	AAT TTG CAG TTT GTT GAA CAA TTT CAA TTA CAA TGT GAA GAT ACG TGT Asn Leu Gln Phe Val Glu Gln Phe Gln Leu Gln Cys Glu Asp Thr Cys 385 390 395 400	1200
40	GAT AAA TCC TAT GTA GAA GTG AAA GCT GAC GCT GAT TTT CGA CCT ACT Asp Lys Ser Tyr Val Glu Val Lys Ala Asp Ala Asp Phe Arg Pro Thr 405 410 415	1248

	GGA TAT CGA TTT TGT TGT TCG CGA GTG CCA CGT CAT ATT TTT CAA TCT Gly Tyr Arg Phe Cys Cys Ser Arg Val Pro Arg His Ile Phe Gln Ser 420 425 430	1296
5	GCG ACA AAC GAG ATG GTA GTA ATA TTT CGC GGT TTT GGT GGT GCG GGA Ala Thr Asn Glu Met Val Val Ile Phe Arg Gly Phe Gly Gly Ala Gly 435 440 445	1344
	AAT GGC TTT AAA GCT AAA ATT TGG TCA AAC GTA GAT GAT GAT ATA GCT Asn Gly Phe Lys Ala Lys Ile Trp Ser Asn Val Asp Asp Asp Ile Ala 450 455 460	1392
10	AAT ACA ATT GTA ACA ACT GAA ATG GCA AAA ATT TCG GAA AAA ATA CCG Asn Thr Ile Val Thr Thr Glu Met Ala Lys Ile Ser Glu Lys Ile Pro 465 470 475 480	1440
15	AAG CTA ACA GTT CCA ATA GTT AAA ACT ATT ACC ACT CCT ACA ATA ACA Lys Leu Thr Val Pro Ile Val Lys Thr Ile Thr Thr Pro Thr Ile Thr 485 490 495	1488
	ACT ACT ACT GCT TTC ATG ATA TCA CCC AAG AAA GGC AAT GTC ACC GCC Thr Thr Ala Phe Met Ile Ser Pro Lys Lys Gly Asn Val Thr Ala 500 505 510	1536
20	ACG AGA GTT GCT ATC ACT ACG CCG ACT ACT ACA ATT ACT ACG ACT Thr Arg Val Ala Ile Thr Thr Thr Pro Thr Thr Thr Ile Thr Thr Thr 515 520 525	1584
	ATT GCC GGT ACG GTA CCA ATC ACC GTA ACT AAT AAT ACT ACC CCT GTA Ile Ala Gly Thr Val Pro Ile Thr Val Thr Asn Asn Thr Thr Pro Val 530 540	1632
25	GTA AGT GAA ACT TTA CCA TCA TTG CCA GTC AAG ATT CGA AAC AAA ATA Val Ser Glu Thr Leu Pro Ser Leu Pro Val Lys Ile Arg Asn Lys Ile 545 550 555 560	1680
30	GGT GCA TGC GAA TGT GGT GAA TGG ACA GAA TGG ACA GGT CCA TGC TCT Gly Ala Cys Glu Cys Gly Glu Trp Thr Glu Trp Thr Gly Pro Cys Ser 565 570 575	1728
	CAA GAA TGT GGC GGT TGC GGA AAA CGT CTT CGA ACA CGT CAG TGT TCA Gln Glu Cys Gly Gly Cys Gly Lys Arg Leu Arg Thr Arg Gln Cys Ser 580 585 590	1776
35	TCA GAT ACG GAA TGT AGA ACA GAA GAA AAA CGT GCG TGT GCT TTT AAA Ser Asp Thr Glu Cys Arg Thr Glu Glu Lys Arg Ala Cys Ala Phe Lys 595 600 605	1824
	GTT TGC CCA TAC GGG ACT AAT TTC CTT ATC AAT AAT GGA GAG TTT CAT Val Cys Pro Tyr Gly Thr Asn Phe Leu Ile Asn Asn Gly Glu Phe His 610 620	1872
40	ATA CTT TGG AAG GGC TGC TGT GTT GGT CTA TTC CGA TCG GGA GAT ATG Ile Leu Trp Lys Gly Cys Cys Val Gly Leu Phe Arg Ser Gly Asp Met 625 630 635 640	1920

		TCA Ser															1968
5		CTG Leu															2016
		AAA Lys															2028
10	(2)	INFO	ORMAT	rion	FOR	SEQ	ID 1	NO:34	1:								
			(i) \$	(A)	LE1	NGTH	: 670 amino	ERIST 6 ami o aci	ino a id		5						
15		(:	Li) N	OLE	CULE	TYPI	E: pi	rote	Ln								
		(2	ki) S	SEQUE	ENCE	DES	CRIP	rion:	SEÇ	Q ID	NO:	34:					
	Gln 1	Ser	Val	Glu	Asn 5	Tyr	Arg	Thr	Gln	Lys 10	Ala	Ser	Arg	Asn	Thr 15	Leu	
20	Asp	His	Ile	Lys 20	Gln	Leu	Ile	Thr	Leu 25	Asn	Val	Gln	Arg	Glu 30	Ile	Gly	
	Asn	Ile	Asp 35	Asp	Lys	Thr	Leu	Ala 40	Asp	Glu	Ile	Val	Leu 45	Gln	Arg	Arg	
	Asp	Pro 50	Glu	Ala	Lys	Trp	His 55	His	Asn	Glu	Leu	Phe 60	Ile	Asn	Asp	Pro	
25	Asp 65	Ala	Tyr	Tyr	Gln	Gly 70	Asp	Val	Asp	Leu	Ser 75	Glu	Lys	Gln	Ala	Glu 80	
	Ile	Leu	Ser	Glu	Hiş 85	Phe	Lys	Asn	Glu	Ile 90	Ala	Leu	Thr	Glu	Lys 95	Asp	
30	Asp	Thr	Ile	Ile 100	Arg	Arg	Lys	Lys	Ser 105	Ile	Gly	Arg	Glu	Pro 110	Phe	Tyr	
	Val	Arg	Trp 115	Asn	His	Lys	Arg	Pro 120	Ile	Ser	Tyr	Glu	Phe 125	Ala	Glu	Ser	
	Ile	Pro 130	Leu	Glu	Thr	Arg	Arg 135	Lys	Ile	Arg	Ser	Ala 140	Ile	Ala	Met	Trp	

Glu Glu Arg Thr Cys Ile Arg Phe Gln Glu Asn Gly Pro Asn Val Asp 145 150 155 160

	Arg	Ile	Val	Phe	Asn 165	Asp	Gly	Gly	Gly	Cys 170	Ser	Ser	Phe	Val	Gly 175	Arg
	Thr	Gly	Gly	Thr 180	Pro	Gly	Ile	Ser	Ile 185	Ser	Thr	Pro	Gly	Cys 190	Asp	Ile
5	Ile	Gly	Ile 195	Ile	Ser	His	Glu	Ile 200	Gly	His	Thr	Leu	Gly 205	Ile	Phe	His
	Glu	Gln 210	Ala	Arg	Arg	Asp	Gln 215	Lys	Asn	His	Ile	Phe 220	Ile	Asn	Tyr	Asn
10	Asn 225	Ile	Pro	Ser	Ser	Arg 230	Trp	Asn	Asn	Phe	Phe 235	Pro	Leu	Ser	Glu	Tyr 240
	Glu	Ala	Asp	Met	Phe 245	Asn	Leu	Pro	Tyr	Asp 250	Thr	Gly	Ser	Val	Met 255	His
	Tyr	Gly	Ser	Tyr 260	Gly	Phe	Ala	Arg	Asn 265	Pro	Tyr	Glu	Pro	Thr 270	Ile	Thr
15	Thr	Arg	Asp 275	Lys	Phe	Gln	Gln	Tyr 280	Thr	Ile	Gly	Gln	Arg 285	Glu	Gly	Pro
	Ser	Phe 290	Leu	Asp	Tyr	Ala	Ser 295	Ile	Asn	Ser	Ala	Tyr 300	Arg	Cys	Thr	Glu
20	Gln 305	Cys	Ala	Asp	Met	His 310	Cys	Asp	His	Asn	Gly 315	Tyr	Pro	Asp	Pro	Asn 320
	Asn	Cys	Ala	Lys	Cys 325	Leu	Cys	Pro	Asp	Gly 330	Phe	Ala	Gly	Arg	Thr 335	Cys
	Gln	Phe	Val	Gln 340	Tyr	Thr	Ser	Cys	Gly 345	Ala	Leu	Ile	Lys	Ala 350	Arg	Lys
25	Met	Pro	Val 355	Thr	Ile	Ser	Ser	Pro 360	Asn	Tyr	Pro	Asn	Phe 365	Phe	Asn	Val
	Gly	Asp 370	Gln	Суѕ	Ile	Trp	Leu 375	Leu	Thr	Ala	Pro	Arg 380	Gly	Gly	Phe	Val
30	Asn 385	Leu	Gln	Phe	Val	Glu 390	Gln	Phe	Gln	Leu	Gln 395	Cys	Glu	Asp	Thr	Cys 400
	Asp	Lys	Ser	Туr	Val 405	Glu	Val	Lys	Ala	Asp 410	Ala	Asp	Phe	Arg	Pro 415	Thr
	Gly	Tyr	Arg	Phe 420	Cys	Cys	Ser	Arg	Val 425	Pro	Arg	His	Ile	Phe 430	Gln	Ser
35	Ala	Thr	Asn 435	Glu	Met	Val	Val	Ile 440	Phe	Arg	Gly	Phe	Gly 445	Gly	Ala	Gly

	Asn	Gly 450	Phe	Lys	Ala	Lys	Ile 455	Trp	Ser	Asn	Val	Asp 460	Asp	Asp	Ile	Ala
	Asn 465	Thr	Ile	Val	Thr	Thr 470	Glu	Met	Ala	Lys	Ile 475	Ser	Glu	Lys	Ile	Pro 480
5	Lys	Leu	Thr	Val	Pro 485	Ile	Val	Lys	Thr	Ile 490	Thr	Thr	Pro	Thr	Ile 495	Thr
	Thr	Thr	Thr	Ala 500	Phe	Met	Ile	Ser	Pro 505	Lys	Lys	Gly	Asn	Val 510	Thr	Ala
10	Thr	Arg	Val 515	Ala	Ile	Thr	Thr	Thr 520	Pro	Thr	Thr	Thr	Ile 525	Thr	Thr	Thr
	Ile	Ala 530	Gly	Thr	Val	Pro	Ile 535	Thr	Val	Thr	Asn	Asn 540	Thr	Thr	Pro	Val
	Val 545	Ser	Glu	Thr	Leu	Pro 550	Ser	Leu	Pro	Val	Lys 555	Ile	Arg	Asn	Lys	Ile 560
15	Gly	Ala	Cys	Glu	Cys 565	Gly	Glu	Trp	Thr	Glu 570	Trp	Thr	Gly	Pro	Cys 575	Ser
	Gln	Glu	Cys	Gly 580	Gly	Cys	Gly	Lys	Arg 585	Leu	Arg	Thr	Arg	Gln 590	Cys	Ser
20	Ser	Asp	Thr 595	Glu	Cys	Arg	Thr	Glu 600	Glu	Lys	Arg	Ala	Cys 605	Ala	Phe	Lys
	Val	Cys 610	Pro	Tyr	Gly	Thr	Asn 615	Phe	Leu	Ile	Asn	Asn 620	Gly	Glu	Phe	His
	Ile 625	Leu	Trp	Lys	Gly	Cys 630	Cys	Val	Gly	Leu	Phe 635	Arg	Ser	Gly	Asp	Met 640
25	Cys	Ser	Ala	Leu	Asp 645	Asp	Asn	Glu	Asn	Pro 650	Phe	Leu	Lys	Phe	Leu 655	Glu
	Ser	Leu	Leu	Asn 660	Met	Gln	Asp	Ser	Arg 665	Lys	Asn	Asp	Asn	Leu 670	Pro	Asp
30	Ser	Lys	Lys 675	Lys												

- (2) INFORMATION FOR SEQ ID NO:35:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA (genomic)

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CATCTCGAGA TCAGTGGAAA ATTATCGAAC G

31

- (2) INFORMATION FOR SEQ ID NO:36:
 - (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 31 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
 - (ii) MOLECULE TYPE: DNA (genomic)
- 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ATTGAATTCA CTTCTTTTTC GAGTCAGGCA A

31

While various embodiments of the present invention have been described in detail, it is apparent that modifications and adaptations of those embodiments will occur to those skilled in the art. It is to be expressly understood, however, that such modifications and adaptations are within the scope of the present invention, as set forth in the following claims: